

공기업 경영 평가 제도의 성과 분석

최성락*·박민정**

〈目 次〉

- I. 서론
- II. 이론적 배경
- III. 분석틀
- IV. 공기업 경영평가 제도의 성과 분석
- V. 결 론

〈요 약〉

평가제도는 단순히 평가 그 자체에 의미가 있는 것이 아니라, 해당 평가 대상자들로 하여금 평가자가 원하는 방향으로 변화를 이끌어내는데 그 목적이 있다. 따라서 평가 제도가 제대로 시행되는가를 검토하기 위해서는 평가 부문에서 개선 효과가 발생하고 있는가를 검토하여야 한다. 본 연구는 한국에서 가장 대표적인 평가 제도의 하나인 공기업 경영평가 제도를 대상으로, 공기업 평가 제도가 원래 의도하였던 평가 항목의 개선 효과를 이끌어내고 있는가를 평균 분산에 대한 가설 검정 및 왜도 첨도 분석 방법을 사용하여 검토하였다.

2001-2005년 사이의 평가 결과를 분석한 결과, 공기업 평가의 각 부문인 종합 경영, 주요사업, 경영관리에서 유의미한 개선이 이루어졌다고 볼 수 없었다. 전체적으로도 공기업 평가제도는 공기업의 평가 점수의 개선을 이끌어내지 못했다. 기타 연구들에서는 공기업의 전반적인 효율성이 향상하고 있는 것으로 나타나고 있는데 평가 제도 점수에서는 유의미한 향상이 이루어지지 않고 있는바, 이는 현재 공기업 평가 제도가 공기업의 효율성 등을 제대로 파악하지 못하고 있으며, 평가 시스템 및 평가 지표 등에 한계가 존재하고 있다는 것을 시사한다.

【주제어: 공기업, 공기업 평가, 메타 평가】

* 동양공전 경영학부 전임강사(haihaba@chol.com).

** 원광대학교 행정학과 전임강사(mjpark@wku.ac.kr).

논문접수일(2008.11.13), 수정일(2009.2.23), 게재확정일(2009.3.5)

I. 서론

1990년대 이후 신공공관리가 확대되면서 정부에 대한 책임 경영이 중시되었다. 이러한 책임 경영 중시에 의하여 공공 부문의 책임성과 효율성을 담보할 수 있는 여러 평가 제도가 대두되게 되었다. 우리나라에서만도 정부 기관평가, 기금 등이 새로 제도화되면서 평가 제도의 중요성이 증가되고 있다. 그런데 과연 이렇게 시행되는 각종 평가 제도들이 정말로 해당 공공 부문의 책임성과 효율성, 생산성 증대에 효과가 있는가를 별도의 문제이다. 평가 제도는 해당 평가 대상자들로 하여금 평가자가 원하는 방향으로 이끌기 위해서 만들어지는 것인데, 만약 평가 제도가 시행되더라도 평가 대상자들의 행동 변화를 이끌어내지 못한다면 평가는 단순히 평가로 끝나는 것이고 효율성 증대라는 원래의 목적을 달성하지 못하게 된다. 따라서 평가 제도는 단순히 평가한다는 그 자체에 의미가 있는 것이 아니라 평가를 해서 해당 평가 대상자의 행동에 실질적으로 변화를 이끌어 냈는가가 의미 있는 것이다.

이와 같이 평가 제도가 실질적으로 목표하는 바를 이끌어냈는가가 중요한 의미를 가짐에도 불구하고 그동안 평가제도의 도입, 제도적 개선점을 논의하는 연구는 다수 존재하였지만, 실제 평가 제도가 원하는 결과를 이끌어냈는가 하는가에 대한 연구는 많지 않았다. 그 이유는 평가 제도와 평가 지표가 지속적으로 변화하여 일관성 있는 연구가 어려운 등 연구의 한계가 존재하기 때문이기도 하고, 평가 제도 자체에 대한 연구의 역사가 오래되지 않았기 때문이기도 하다.

따라서 본 연구는 한국에서 가장 대표적인 평가 제도의 하나인 공기업 경영평가 제도를 대상으로, 공기업 평가 제도가 원래 의도하였던 평가 항목의 개선 효과를 이끌어내고 있는가를 살펴보았다. 개선 효과 판단은 기존 공기업 평가 결과를 바탕으로 평균, 분산에 대한 가설 검정, 그리고 왜도, 첨도 등 기술적 통계량을 분석하는 방법을 활용하였다. 방법론상으로 이러한 지표를 통하여 어느 정도 개선 효과 여부를 판단할 수 있는가 하는 한계는 존재하지만, 아직까지 평가 결과를 바탕으로 개선 효과를 파악한 연구가 많지 않다는 점, 그리고 아직까지 평가 결과가 누적된 것이 적고 해당 평가 기준과 지표가 많이 변하는 현실적 한계에서는 의미를 가질 수 있을 것이다.

II. 이론적 배경

1. 평가 제도에 대한 평가

어떤 대상에 대한 검사 및 평가가 그 검사 및 평가 대상에 대해 미치는 영향

은 지대하다. 개인에 대한 평가 점수는 해당 개인의 입사, 고용, 승진 등에 많은 영향을 미치며, 조직에 대한 평가도 그 조직의 운영 및 방향, 조직원의 사기 등에 큰 영향을 미친다(성태제, 2003). 특히 평가는 과거의 잘된 점과 잘못된 점에 대한 정보만을 주는 것이 아니라, 앞으로의 행위에 영향을 미친다는 점에서 보다 큰 의미를 지니고 있다. 평가의 내용과 방향은 평가 대상의 행위를 구속할 수 있으며, 향후의 방향을 특정 지을 수 있다. 평가자는 평가 내용과 측정 방법을 변화시킴으로서 평가 대상자의 행동을 평가자가 의도하는 방향으로 변화시킬 수 있다(성태제, 2003).

평가는 개인과 기업, 정부 등 모든 조직에서 항상 이루어져 왔다. 그런데 정부 부문에서는 특히 1990년대 이후 신공공관리론이 대두되면서 정부 업무 평가가 보다 고도화되고 정밀화되는 방향으로 진전되게 되었다. 신공공관리론에서는 정부의 책임 경영을 중시하고 기업과 동일한 정도의 효율적 운영, 생산성 확보 등을 요구하는데, 정부의 책임성, 효율성, 생산성 등을 담보하기 위한 중요한 제도적 맥락으로 등장한 것이 바로 새로운 평가 시스템이었다. 이러한 평가에 대한 중요성이 증가하면서 한국에서는 2000년 ‘정부업무 등의 평가에 대한 기본법’이 제정되어 정부 기관 평가에 대한 새로운 장이 형성되었다. 그리고 정부 부처 외에 1999년부터는 정부출연 연구기관 평가와 책임운영기관 평가가 이루어졌으며, 2005년부터는 정부산하기관 경영평가가 시작되었다. 또한 1984년 이후 계속 되어온 정부투자기관 경영평가도 2000년 이후 보다 정밀한 평가 방법을 도입하였으며, 이외에 기금 평가, 학교 기관 평가 등이 시작되었다.

그러나 위와 같은 평가 제도가 도입된다고 하여 당연히 정책 및 업무 개선이 이루어지는 것은 아니다(Schneider, 1986). 평가가 적정히 이루어지고, 평가 결과의 피드백이 원활히 이루어지며, 평가 결과가 적절하게 활용될 경우에 평가는 의미가 있다. 그러나 현대 국가에서 수행하는 다양한 형태의 정책 평가가 현실의 정책 과정을 개선하기에는 평가 결과의 질이 너무 낮다는 비판이 제기되었고(임성욱, 2003), 이러한 정책 평가의 유용성에 대한 의문에 대응하기 위하여 평가 활동 자체의 적정성과 합리성을 다시 평가하고자 하는 ‘평가에 대한 평가’에 대한 논의 필요성이 중시되었다(이찬구, 2003).

‘평가에 대하여 평가’하는 것은 일반적으로 메타 평가(meta evaluation)라고 한다(Sciven, 1981). 그런데 메타 평가는 ‘평가에 대하여 평가’하는 것이 구체적으로 어떤 것인가에 대한 견해 차이에 따라서 크게 3가지로 구분된다(문영세, 1996).

첫째로 ‘평가’를 제일 광범위하게 규정하는 것으로서 ‘평가 체제 전체’를 평가하고자 하는 것이 메타 평가라는 견해가 있다(Larson & Berliner, 1983; 김명수, 1993). 이러한 입장에 있는 대표적인 학자로는 Larson & Berliner 등이 있는데, Larson & Berliner(1983)는 메타평가를 평가투입, 평가과정, 평가결과 등을 포

함하는 평가체계 및 절차에 대한 평가로 이해하였다. 즉 단순히 평가 결과만을 평가 하지 않고, 총괄적인 평가 시스템 및 평가 제도 전반을 평가하고자 하는 것이다. 현재 우리나라에서 가장 일반적으로 연구되는 메타 평가는 바로 이 광의의 개념이다. 박종수(2003)는 메타평가를 평가맥락, 평가투입, 평가과정, 평가결과 및 평가활용 등을 포함하여 평가 체제론적 관점에서 평가하였고, 황병상·강근복(2005)은 정부출연연구기관 평가를 대상으로 평가기획, 평가수행, 그리고 평가결과 및 활용으로 나누어 메타평가를 실시하였다. 문원식(2005)은 지방행정 혁신평가에 대하여 평가목적, 접근방법, 평가체계(평가주체, 평가대상, 평가방법), 평가결과의 환류 등 4가지를 기준으로 미국 제도와 비교분석 하였으며, 공병천(2005)은 중앙행정기관평가를 중심으로 평가기준, 평가지표, 평가내용, 평가주체, 평가일정을 기준으로 메타평가를 실시하였다. 이와 같이 평가 시스템 전반에 대한 평가가 이루어지는 것이 첫 번째 견해의 메타 평가로 볼 수 있다.

두 번째로는 메타 평가를 중범위로 정의하는 것으로 평가 자체에 대한 평가를 의미한다(Cook & Gruder, 1978). 기존 평가와 다른 조사 설계 등을 이용해서 기존의 평가 결과가 제대로 된 것인지를 평가하는 경험적인 총괄 평가로서의 성격을 지니고 있다. 기존의 공식적인 평가 방법 외에 BSC, DEA 등 새로운 방법론을 적용하여 평가를 하거나(이환범 외, 2005; 이형우 외, 2005), 새로운 평가지표를 제시하는 등의 연구(이석원, 2005)가 이러한 중범위의 메타평가에 속한다고 볼 수 있다.

마지막으로 가장 협의로 사용되는 메타 평가는 평가 종합(evaluation synthesis)이다(Chelimsky, 1985; 노화준, 1995). 평가 종합은 동일한 프로그램이나 정책에 대한 여러 평가연구 결과를 체계적으로 종합한 것으로서, 하나 혹은 여러 개의 기존 평가로부터 얻어진 결과를 재분석하는 평가형태이다. 이 평가 종합에서는 ‘정책이나 프로그램의 평균적인 효과는 무엇인가’, ‘해당 정책은 어디 그리고 누구에게 특히 효과가 있는가’ 등에 대한 의문에 답하기 위해서 주로 활용된다(황혜신, 2006). 평가 종합 측면에서 이루어진 대표적인 연구로는 황혜신(2006)의 연구가 있다. 평가 종합론에 Fischer의 정책평가 논리를 접목하여 국민연금 정책에서 각 수준별로 어떠한 평가연구들이 있었는지를 살펴보았다.

2. 공기업 평가 제도의 평가에 대한 선행 연구

현재 한국에서 시행되고 있는 공공부문에 대한 평가 제도로는 정부 기관 평가, 정부 출연연구기관 평가, 책임운영기관 평가, 정부산하기관 경영평가, 기금평가, 정부투자기관 경영평가 등이 있다. 이 중에서 가장 오랜 역사를 가지고 가장 발달된 형태의 평가 제도를 보유하고 시행되는 평가 부문은 정부투자기관 경영평가이다. 정부 공기업 부문에 대해 평가를 시행하는 정부투자기관 경영평

가는 1984년부터 현재까지 시행되고 있는데, 그동안 많은 시행착오를 거치면서 한국의 대표적인 평가 제도로 정착되었다. 이러한 오랜 역사를 가진 공기업 평가 제도의 성과 등에 대하여 평가를 한 연구도 많이 이루어졌는데, 최근의 대표적인 선행연구의 주요 내용과 결과를 살펴보면 다음과 같다.

1) 공기업 경영평가 제도 및 개선방안에 대한 연구

우선 현행 공기업 평가제도의 도입과 관련한 배경과 목적 등에 대한 논의를 한 연구들을 살펴보면 송대회(2003)는 1984년부터 시행된 우리나라 정부투자기관 경영평가제도의 구축 배경과 그 동안의 변천과정, 그리고 대내외적 파급효과를 분석하였다. 신열(2004)은 공기업 경영평가의 목적을 제시하고 있었으며, 노화준(2001)은 공기업 경영평가가 정치적 책임성 확보와 지지, 성과개선을 위한 관리적 수단으로서 정책평가적 의미를 지닌다고 보았다. 김종희(2006)는 정부에 대한 평가도 중요한 관리과정 중의 하나이기에 공기업에 대한 평가도 이와 같다고 보아, 공기업에 대한 평가가 필요함을 제시하였다.

이러한 공기업 평가제도 배경에 관한 연구 외에 공기업 경영평가와 관련한 선행 연구는 평가제도 시행과 관련하여 나타나는 실질적인 문제점 및 이를 개선하는 방안을 제안하는 연구들이 많은 비중을 차지하고 있다. 오연천·곽채기(2003)는 경영평가방법 및 1년 단위 평가주기상의 문제점, 경영평가단 구성 및 운영상의 문제점, 평가단의 전문성과 평가의 계속성 및 일관성 확보, 경영평가 결과의 활용 및 환류기능 미흡 등을 경영평가 제도의 주된 문제점으로 지적하였다. 김준기(2001)는 현행 정부투자기관 경영평가제도의 성공적 운영을 위한 전제조건인 경영자율성 보장을 주된 문제로 제기하면서, 평가기간과 시기, 지표체계의 중복성과 복잡성, 기관특성을 반영 못하는 지표체계, 계량지표의 증가와 비계량의 감소 요구 증대 및 두 지표간의 연계성 빈약, 공기업 자체 통제 불가능 지표 존재, 피드백의 내실화 부족 등을 제시한다. 윤성식 외(1999)는 정부투자기관의 경영평가에 대한 실증적인 조사로 경영평가제도의 문제점을 지적하고, 그 수용도의 낮은 원인을 경영평가에 대한 평가대상기관의 인식, 평가기관, 평가지표로 보고 있다. 이규환(2000)의 경우도 경영평가제도자체의 문제점을 지적하였는데, 경영평가 전담기구의 부재와 인력·예산의 부족, 경영평가 중용성에 대한 미인지, 평가대상기관의 경영혁신요구를 반영하지 못하는 지표 등을 문제점으로 삼았다.

신열·오철호(2002)의 연구에서는 경영평가가 발전하기 위한 가장 중요한 과제를 평가체제, 평가지표, 평가방법에 대해 전반적으로 관리하는 평가주체에 대한 개선에서부터 접근해야 할 것으로 지적하였다. 특히 평가지표의 체계, 구성은 맹목적 계량화보다는 평가대상기관에 따른 차별화를, 평가방법상의 문제들

해결하기 위해서는 개별 평가대상기관의 특수성과 일반성을 조화시키는 노력이 필요하다고 제언하고 있다. 광채기(2003)의 연구에서는 13개 공기업을 대상으로 경영혁신의 촉진, 경영투명성, 고객만족도, 경영효율성, 조직생산성, 공기업관리 정책개선 등에 걸쳐 경영평가제도의 효과를 실증적으로 분석하였다. 경영실적평가를 통한 환류기능이 어떻게, 그리고 어느 정도 효과적으로 수행되어 왔는가를 하는 점을 지난 20년 동안의 경영실적 평가 결과를 기록한 경영실적평가보고서에 대한 내용분석 및 설문조사를 통해 확인하였다. 정부투자기관 경영평가제도가 실제 공기업의 비효율성을 얼마나 효과적으로 통제하여왔는가를 경험적으로 분석 평가하였는데, 분석 결과 경영평가제도는 정부투자기관의 경영관리 도구로서 경영 목표 설정, 경영진단 및 환류 기능을 비교적 성공적으로 수행한 것으로 보았다. 경영평가제도는 정부투자기관의 경영혁신 촉매 및 확산, 경영투명성 제고, 고객만족도 제고, 경영효율성 제고 등의 측면에서 긍정적인 성과를 창출하였으며, 한계로서는 정부투자기관의 경영자율성을 확대하는데 있어서는 성공적이지 못하였던 점이 지적되었다. 장지인·광채기(2000)는 사장경영계약 이행실적 평가 과정에서 사장의 개인적인 경영노력의 결과, 즉 사장 개인의 경영성과와 기관 차원의 경영성과를 엄격하게 구별하는 것이 현실적으로 어렵다고 보았으며, 평가제도가 이원화됨으로서 관련 경영목표 관리 체계에 혼선이 초래될 수 있는 문제점이 있다고 보았다.

이러한 경영평가에 대한 문제점을 지적하는 등의 부정적인 시각뿐 아니라, 실제로 경영평가제도의 긍정적인 효과를 연구한 문헌들도 있다. 송대희(2003)는 공기업 임직원에 대한 설문분석 방법을 활용하여, 정부투자기관 경영평가 제도가 공기업의 경영 효율성을 개선하고 경영 혁신 의지를 확산하는데 긍정적인 효과를 가져온 것으로 평가하였다. 박석희(2006)는 1965년부터 2004년까지 13개 정부투자기관에 대한 패널 데이터를 활용하여 정부투자기관 경영평가제도 도입 및 개선이 공기업들의 관리업무 비율, 인건비 비율, 금융 비율을, 인원증감률 등을 감소시켜 전반적으로 조직 역량이 증가되었다는 결과를 도출하였다.

그리고 개선방안을 제시함에 있어 궁극적으로 공기업의 책임경영을 위한 논의들을 한 연구들도 있는 것으로 이상철·성도경(2002)은 공기업의 책임경영을 위한 성공요인들을 분석한 바 있고, 송대희·송명희(1994)는 공기업의 책임경영 운영에 대한 평가 및 방향을 제시해주었다. 이외에도 경영평가제도 자체의 현황과 문제점을 제기하면서 발전방향을 제시한 연구로 오연천(1996), 배용수(2006), 신열(2002; 2004), 김종희(2006)의 연구가 있다.

2) 공기업 성과평가 지표에 대한 연구

보다 미시적인 접근으로 공기업 성과평가 지표 자체에 대해서 분석적이며,

구체적으로 연구한 문헌들로는 DEA와 서베이 방법을 이용한 새로운 평가지표를 적용하면서 현행 공기업의 성과지표와 평가방법을 비판적으로 검토하고자 한 이환범·송건섭·김병문(2005)의 연구가 있으며, 이석원(2005)의 연구에서는 기존의 성과지표의 한계를 지적하면서 정부투자기관경영평가에 있어서 조정성과지표를 사용하여 공공기관의 성과를 측정에 적용할 것을 제안 하고 있다. 김중희(2007)는 지방공기업 경영평가에 있어서 그동안 비계량지표의 가중치가 어떻게 조정되었으며, 평가결과 순위가 공기업의 규모와 어떤 상관관계가 있고, 순위가 고착화되고 있는 원인 등에 대한 분석과 평가결과에 대한 공기업의 수용도 제고방안을 모색하고 있다. 또한 나영·이창우(2001)는 서울특별시의 6개 지방공사, 공단의 경영평가결과를 토대로 계량지표와 비계량지표의 선정 및 적용의 적정성과 경영평가지표의 사전적 내부활용도 등을 검증하여 보았다. 백승천(2003)은 현행 지방공기업 경영평가지표가 단기적 측면에서 과거에 초점을 둔 평가지표가 많고, 후행지표의 동인이 되는 선행지표가 부족하여 지방공기업 자체의 특성을 반영하지 않는 가중치 설정이 문제이며, 평가지표 수가 지나치게 많다고 지적하고 있다. 이종수(2003)의 연구에서도 지방 공기업 경영평가제도의 핵심적 개선사항은 평가지표의 개선에 있는 것으로 보고 공기업의 경쟁력 제고를 통한 공기업 경영평가 전략 탐색의 이론적 토대 구축 방안의 하나로서 조직문화지표개발에 초점을 두었다. 특히 공기업 평가시 문화지표 개발의 필요성을 제기하면서, 이의 도입으로 현 지표의 문제점 등을 개선할 것으로 제안하고 있다. 박영범(1998)은 경영환경 변화를 고려하여 비계량지표 개선을 해야한다는 제안을 하였고, 박명호·이성욱(2000)도 역시 비계량지표상의 제반개선사항을 논의하였다. 이외에도 지표상의 개선을 지적한 논문으로 장지인·이정구(1998), 경영평가지표의 인과관계에 대한 연구를 한 광병환(2005), BSC 관점에서 경영평가 지표 개선에 관한 연구를 한 이형우·임동환(2005) 등이 있다.

3. 선행연구의 한계와 본 연구의 접근법

1) 선행연구의 한계

공기업 평가제도 평가에 대한 위의 선행 연구들은 명시적으로 메타 평가를 명시하지는 않았지만, 그 내용상 실질적으로는 메타 평가의 일환으로 볼 수 있다. 장지인·곽채기(2000), 오연천·곽채기(2003) 등의 연구는 공기업 평가 제도 전반에 관한 사항을 검토하고 개선 방안을 논의한 것으로 광의의 메타 평가의 일환으로 볼 수 있으며, 곽채기(2003)의 연구는 지난 20년간의 평가 결과를 대상으로 2차적 분석을 한 것으로 평가종합적 성격을 지니고 있다. 그리고 박석희(2006)의 연구는 공기업 평가제도가 실질적으로 조직 역량에 어떠한 영향을 미쳤는가를 평가하고자 하는 것으로 광의의 메타 평가에 속한다고 판단할 수 있

다. 그리고 공기업 성과평가 지표에 대한 연구들은 중범위의 메타평가에 속한다.

이러한 선행연구들의 한계를 살펴보면, 우선 대부분의 공기업 경영평가 제도의 개선을 논의하는 연구가 이론적이며 규범적인 논의에 머물고 있다는 점이다(박석희, 2006). 곽채기(2003), 박석희(2006)의 연구가 실증적인 자료를 바탕으로 공기업 평가제도에 대한 평가를 시행하고 있으며, 나머지 대부분의 연구는 규범적이고 이론적인 바탕 위에서 평가제도에 대한 논의가 이루어지고 있다. 그리고 실증 분석의 경우에도 대부분 설문 분석을 중심으로 하고 있어 평가제도의 실질적인 효율성 및 생산성 증대 효과를 파악하기 어려운 한계를 지니고 있다.

그리고 곽채기(2003), 박석희(2006) 등은 실증적인 자료를 바탕으로 평가제도에 관한 평가를 수행하고 있지만, 이들 평가는 평가 제도를 직접적으로 평가 대상으로 하고 목적으로 삼고 있는 사항을 측정하는 것이 아니라 경영 투명성 제고, 조직 역량 등 기타 사항을 중심으로 평가 제도의 성과를 파악하고 있다는 한계점을 가진다. 평가 제도는 자체적으로 추구하는 가치가 있고, 공기업을 어느 방향으로 이끌고자 하는 목적을 지니고 있다. 평가 제도가 제대로 기능한다는 것은 공기업이 바로 이런 평가 제도의 가치와 방향에 이끌려가고 있다는 것을 말하는 것이며, 바로 그러한 가치와 방향 측면에서 진전이 있는가가 평가 제도의 적합성을 판단하는 기준이 되어야 한다. 그러나 곽채기(2003), 박석희(2006) 등의 연구는 조직 역량, 경영 투명성 제고 등 평가 제도가 직접적으로 평가 대상으로 하지 않고 있는 사항을 중심으로 평가하고 있다. 물론 이러한 사항도 공기업의 경쟁력 강화에 큰 기여를 하고 간접적으로 공기업 평가 제도의 목적에 부합하는 사항이기는 하지만, 평가 제도의 직접적 기여도를 평가하기 위해서는 보다 직접적으로 평가 제도가 평가 대상으로 삼는 부문에서 성과가 증진되었는가를 판단할 필요가 있다.¹⁾

2) 본 연구의 접근법

위와 같은 선행 연구의 한계를 보완하기 위하여 본 연구에서는 평가 제도가 평가 대상으로 삼고 있는 부문에서 실질적으로 발전 및 개선이 이루어지고 있는가를 실증적으로 살펴보고자 한다. 2001년부터 2005년 사이의 공기업 평가 결과를 바탕으로, 공기업 평가 제도가 직접 평가하고 있는 부문에서 개선이 있었

1) 중고등학교에서 영어, 수학 등의 시험을 보는 것은 1차적으로 영어 및 수학 실력을 증진시키고, 궁극적으로는 학생의 총체적인 지적 능력을 향상시키기 위함이다. 영어, 수학 시험의 효과를 측정하기 위해서는 ‘학생의 총체적 지적 능력’ 향상을 파악하는 것도 중요하긴 하지만, 이는 [영어 및 수학 실력=>지적 능력]이라는 인과관계가 성립하는가란 문제가 제기된다. 시험의 효과를 보다 본질적으로 측정하기 위해서는 이러한 시험에 의하여 시험의 직접적인 측정 대상인 영어 실력 및 수학 실력이 얼마나 향상되었는가를 파악하여야 한다.

다고 판단할 수 있는가를 기술통계적 방법을 활용하여 검토한다. 이 연구는 기존의 평가 결과들을 2차적으로 분석하여 새로운 시사점을 도출해낸다는 점에서 메타 평가 중 평가 종합 연구의 일환으로 볼 수 있다. 그리고 해당 평가 결과에서 개선 여부를 판단하는 방법으로는 기술적 통계의 일종인 평균, 분산의 변화, 그리고 왜도와 첨도의 변화 추이를 살펴보는 방법을 활용한다.²⁾

Ⅲ. 분석틀

1. 분석 대상

본 연구에서는 공기업 평가 제도가 평가 목적으로 하고 있는 부문에 대하여 실질적으로 개선 효과를 시현하고 있는가를 검토하는 것을 목적으로 한다. 따라서 분석 대상은 공기업 평가 제도에 의한 평가 결과이다.³⁾ 그런데 한국에서 공기업 평가 제도는 그동안 지속적으로 변화되어 왔기 때문에 단순히 평가 결과의 추세를 보고 개선 효과 여부를 판단할 수는 없다. 한국의 공기업 평가는 1984년부터 시행되었는데, 1984년부터 현재까지 공기업 경영평가 제도는 크게 3시기로 구분된다. 1기는 1984년부터 1991년까지로 종합지표, 설립목적 수행지표, 경영관리 지표 등 3개 지표로 크게 구분되었으며, 수, 우, 미, 양, 가의 5개 평가 등급이 있었다. 제 2기는 1992년부터 1998년까지 종합경영부문, 주요사업부문, 관리효율부문, 경영관리 부문 등 4개 부문으로 구분하고, 1996년부터 5등급 평가에서 9등급 평가로 확대되고, 1998년도부터는 평가지표 점수 스케일이 75-100점에서 0-100점으로 확대되었다.

1999년부터의 제 3기는 정부투자기관관리기본법 개정을 통해 객관성 및 공정성을 확대하는 방향으로 경영평가제도가 변화되었다. 특히 사장평가제도가 채택되는 등 책임경영체제가 도입된 시기이다. 2000년도에서는 종합경영부문, 기업성 부문, 공익성 부문, 경영혁신 부문 등 4개 평가 부문을 적용하고 150점 만점 체제였으나, 2001년부터는 종합경영부문, 경영목표 부문, 경영혁신부문 등 3개

2) 본 기술적 통계 지표로서 개선 여부를 판단하는 방법은 이하 분석 방법에서 상술한다.

3) 평가결과를 대상으로 분석을 하기 위해서는 먼저 평가 결과 점수에 대한 신뢰성이 인정되어야만 한다. 많은 평가의 경우, 상대적인 서열을 중시하는 평가, 정치적 요소를 고려한 평가 등으로 해서 평가 결과 점수에 신뢰성이 적은 경우가 존재한다. 하지만 공기업 평가의 경우 공공기관 평가 중 가장 오래되고 대표적인 경영 평가로서, 그동안 지속적인 평가 체계의 개선이 있어왔고, 기업적 성격이 강하여 경영 지표 위주의 평가가 이루어졌기 때문에 기관 평가 중에서는 그 결과의 신뢰성이 가장 인정되는 평가라 할 수 있다.

부문으로 구성하고 100점 만점의 평가 체제를 수립하여 시행하고 있다.

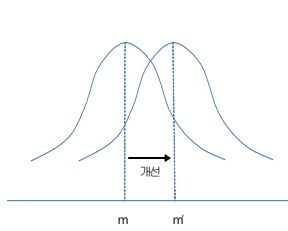
본 연구에서는 공기업 평가 결과를 바탕으로 해당 부문에 대한 개선 효과가 존재하는가를 살펴보고자 하는 것으로, 본 연구가 실행되기 위해서는 평가 지표 부문에 일관성이 있을 것을 전제 조건으로 한다. 그러나 한국 공기업의 경영평가제도는 그동안 지표, 점수 배분 방법, 평가 부문이 지속적으로 변화되어 왔기에 모든 기간 동안 일괄적으로 평가 결과를 바탕으로 분석하기 어려운 측면이 존재한다. 그런데 현재의 공기업 평가 제도는 2001년부터 종합경영 부문, 경영목표 부문, 경영혁신 부문 등 3개 부문으로 안정화되고, 각 부문에 대한 가중치도 종합 경영 30, 주요 사업부문 35, 경영관리 35로 일관성을 지니고 있다. 따라서 본 연구에서는 2001년부터 2005년도까지 5개년을 주요 분석 대상으로 하고자 한다.⁴⁾ 그런데 동 기간 중 2004년도의 경우 가중치가 종합경영 부문 30, 주요사업 부문 40, 경영관리 부문 30으로 변동이 있었다. 2005년도에는 다시 종합경영 30, 주요사업 35, 경영관리 35로 환원하였는바, 분석의 일관성을 위하여 2004년을 제외하고 2001년, 2002년, 2003년, 2005년의 4개년도의 평가 결과를 바탕으로, 해당 부문의 개선 여부를 분석한다.

2. 분석 방법

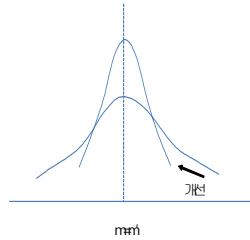
평가 제도는 평가 대상자가 평가하는 부문에 좀 더 초점을 맞추도록 평가 대상자의 행위를 유도한다. 따라서 평가 제도가 제대로 기능하고 있는가를 판단하기 위해서는 평가 지표 부문에 지속적인 개선이 발생하고 있는가를 검토할 필요가 있다. 그리고 지속적인 개선이 발생하고 있는가를 파악하는 가장 일반적인 통계 지표는 평균과 분산이다. 해당 부문에 대한 전체적인 수준이 증가한다면 평균이 상승하게 된다. 평균이 상승한다면 전체적인 수준이 증가하고 있다고 1차적으로 판단할 수 있다. 그리고 같은 평균 하에서라면 분포가 퍼져있는 것보다는 평균 주위에 몰려 있는 것이 더 안정적인 수준 상승을 시현하고 있다고 볼 수 있다. 즉 평균이 동일하다면 분산이 적은 것이 더 높은 수준이라 할 수 있으므로, 평균의 상승과 분산의 감소가 개선을 파악하는 가장 1차적인 지표가 된다.

4) 1기 기간 동안에는 수, 우, 미, 양, 가의 5단계 평가였기 때문에, 본 연구의 평균, 분산, 왜도, 첨도의 분석 방법을 적용하기에는 한계가 있다. 또한 2기(1992-1998)의 기간 동안에는 1995년까지는 5단계 평가였고, 1998년에는 평가점수 범위가 달라지기 때문에 본 분석방법을 적용하기에는 한계가 있다. 따라서 본 연구에서는 3기를 주요 분석 대상으로 하였다.

<그림 1> 평균의 상승과 성과 개선



<그림 2> 분산의 감소와 성과 개선



그러나 평균 점수가 좌우로 넓게 퍼지지 못하고 더 이상 나아갈 수 없는 벽이 존재할 경우에는 평균과 분산으로 개선 여부를 판단하는 것은 한계가 있다. 공기업 평가 점수의 경우 최고점이 100점으로 아무리 개선이 이루어진다 해도 100점을 넘을 수는 없다. 이런 경우 100점 가까이 갈수록 점차 평균의 상승이 어려워지고, 따라서 평균의 상승 여부로 개선 여부를 판단하기 힘들다. 특히 평균이 이미 한계치에 가까운 경우에는 평균의 상승과 하락여부로 개선과 하락을 판단하는 것은 한계가 존재한다. 이와 같이 벽이 존재하여 평균 점수가 어느 한도 이상 증가할 수 없는 경우에는 왜도와 첨도의 변화로서 그 개선 여부를 판단할 수 있다.⁵⁾

먼저 벽이 존재하여 평균이 크게 증가하지 못하는 경우라 하더라도 평가 대상자들의 평가 결과가 개선되는 방향으로 집중된다면 이때는 평가로 인한 개선이 이루어지고 있다고 판단할 수 있는데, 이는 왜도의 변화로 분석할 수 있다.

왜도는 $S_k = \frac{\text{평균} - \text{최빈치}}{\text{표준편차}}$ 로 정의되는데, 이 정의에서는 최빈치가 정의되어야

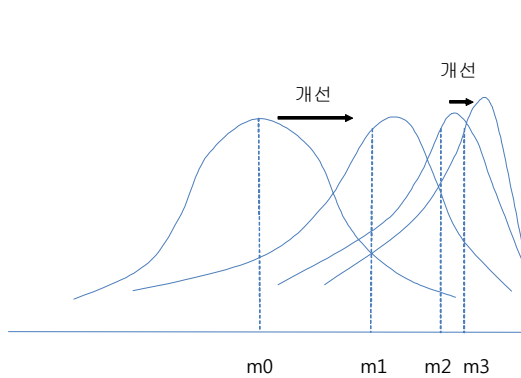
하기 때문에 2차 적률과 3차 적률을 이용한 $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$ 로 일반적으로 사용되고 있

다. 왜도는 분포의 비대칭성을 측정하는 개념으로서, 분포가 오른쪽으로 왜곡된 경우(봉우리가 왼쪽에 나타나는 경우) 왜도가 양의 값을 지니며, 분포가 왼쪽으

5) 더 이상 넘어설 수 없는 벽이 있는 상태에서는 평균보다 왜도와 첨도를 측정하는 것이 성과평가에 더 적절할 수 있다는 견해를 제시한 것은 굴드(Gould)이다. 굴드는 현대 야구 선수의 타율이 4.0을 넘어서지 못하는 것은 야구선수의 실력이 증가하지 않아서가 아니라, 타율 4.0이 인간 능력의 벽이라고 보았다. 타율 4.0이 벽인 상태에서 평균은 더 이상 유의미하게 증가하지 않고 있으나, 왜도와 첨도가 계속 증가하고 있는 것을 보여줌으로서 현대 야구의 수준이 계속 상승하고 있다는 것을 제시하였다. 생물 종의 진화와 관련해서도, 박테리아라는 하향 벽이 존재하는 상태에서 평균은 비록 증가하지만, 왜도와 첨도가 박테리아 수준에서 더욱 집중하는 것을 볼 때 생물 종들이 전체적으로 발전하는 것은 아니라는 주장을 하였다.(Gould, 1996)

로 왜곡된 경우(봉우리가 오른쪽에 나타나는 경우) 왜도는 음이다. 정규분포일 때 왜도는 0이 되며, 그 절대값이 클수록 왜곡 정도가 심해진다. 이러한 왜도의 특징은 벽이 존재할 경우 성과 개선 여부를 판단할 수 있는 시사점을 준다. 벽이 존재하는 경우 비록 개선이 발생한다 하더라도 평균이 더 이상 증가하기는 힘들지만, 분포 자체는 벽이 있는 쪽으로 치우치게 된다. 따라서 100점 만점인 분포에서 평균의 증가가 거의 정체 상태라 하더라도 왜도의 음수 값이 점차 적어진다면 그 성과는 개선되고 있다고 판단할 수 있다(이종욱, 2003).

<그림 3> 왜도의 감소와 성과 개선

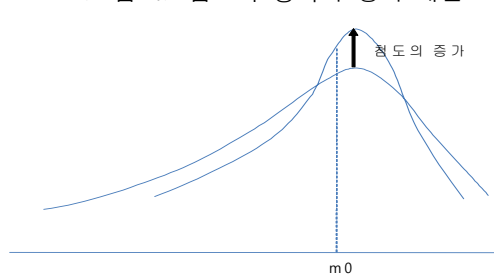


* 벽이 없는 경우에는 평균의 변화(m0→m1)에 의하여 개선 평가.
 벽이 존재하는 경우에는 평균의 변화(m2→m3) 보다는 왜도로 개선 평가.

그리고 분포가 벽에 가까워지는 경우에는 분산보다는 첨도가 성과 개선을 판단하는데 더 유용할 수 있다. 첨도는 분포의 뾰족한 정도로서 $\gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\mu_2^2} - 3$

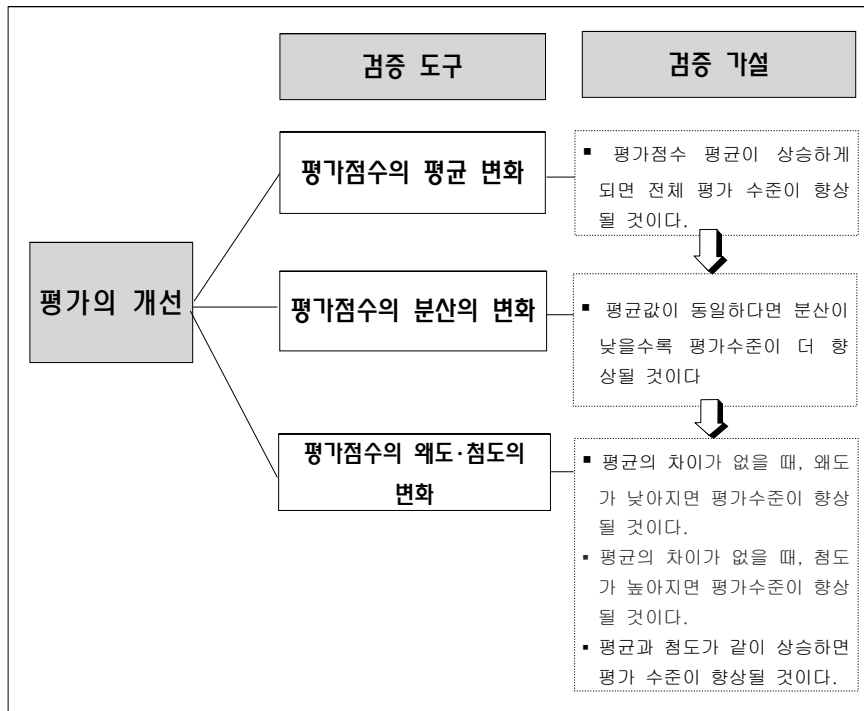
로 정의된다. 정규분포에서 첨도는 0이 되며, 정규분포보다 평균에 집중되는 정도가 높아 분포가 뾰족해진다면 첨도의 값은 양이 되고, 정규분포보다 분포가 평평하다면 첨도의 값은 음이 된다. 즉 첨도의 값이 클수록 평균으로 수렴하는 경향이 강해지고, 이를 기반으로 개선 여부를 판단할 수 있다. 첨도로 개선 여부를 판단할 때는 평균의 변화와 같이 고려해야 한다. 평균이 동일할 때 첨도가 높다면 개선이 이루어진 것으로 판단할 수 있으며, 평균이 상승하고 첨도가 동일할 경우에도 개선이라고 판단할 수 있다. 그리고 평균과 첨도가 같이 상승하면 크게 개선이 이루어졌다고 판단할 수 있다(이종욱, 2003).

〈그림 4〉 첨도의 증가와 성과 개선



본 연구에서는 2001년부터 2005년까지의 공기업 평가 결과를 바탕으로 각 부문의 평균, 분산에 대한 가설 검정, 그리고 각 부문에서의 평균, 분산과 왜도, 첨도의 관계를 살펴봄으로써 평가 부문에 대한 개선이 발생하였는가를 평가한다.6)

〈그림 5〉 연구의 분석틀



6) 본 연구에서는 공기업 경영평가 점수를 대상으로 평균, 분산, 왜도, 첨도 분석을 시행한다. 공기업의 실제 매출, 이익율, 생산성 등 직접적인 자료를 대상으로도 분석이 가능하겠지만, 본 연구는 메타 평가의 일환으로 평가제도 자체를 평가하고자 하는 것이기에 평가 점수를 그 분석 대상으로 한다.

IV. 공기업 경영평가 제도의 성과 분석

1. 평균 분산 검정

공기업 평가는 해당 지표에 대한 평가를 수행함으로써 궁극적으로 공기업이 해당 지표에 대한 부문을 개선시키도록 하는 것을 목적으로 한다. 따라서 공기업 평가 제도가 진정으로 공기업에 영향을 미친다면, 평가 지표 부문이 지속적으로 개선되어야만 한다. 만약 다른 경영 효율성 측면에서 개선이 이루어졌다 하더라도, 막상 평가 지표 부문이 개선되지 않았다면 평가 제도가 의미 있게 운영되고 있다고 보기 힘들다. 따라서 이하에서는 먼저 평균과 분산에 대한 가설 검정을 활용하여 연도별 공기업 평가 결과에 지속적인 개선 효과가 존재하는지, 즉, 평가 결과 사이에 변화가 존재한다고 판단할 수 있는가를 검정한다.

평가 제도가 행위자의 행동에 영향을 미치는 것은 단기간 동안에 발생할 수도 있고, 단기간이 아니라 장기간 동안 영향을 미쳐 변화를 이끌어낼 수도 있다. 따라서 여기서는 단기간과 여러 해 동안의 변화를 평균 검정 및 분산 검정을 통하여 각 기간간 변화가 존재하는가를 살펴본다. 자료가 2001년부터 2005년까지의 기간이기 때문에 1년차 기간으로는 2001년과 2002년, 2002년과 2003년을 살펴보고, 2년차 기간으로는 2001년과 2003년, 2003년과 2005년을 살펴본다. 3년차 기간으로는 2002년과 2005년이 설정될 수 있으며, 4년차 기간으로는 2001년과 2005년을 비교할 수 있다.

먼저 전체 평가 결과에 차이가 있는가를 살펴보면 다음의 <표 1>과 같다.

<표 1> 전체의 평균 및 분산 변화 검정

기간	가 설	평균 검정		분산 검정		
		H0:차이없음 H1:차이있음	P 값	H0:차이없음 H1:차이있음	F 값	P 값
1년간	2001년-2002년 비교	H0 채택	0.567	H0 채택	0.890	0.354
	2002년-2003년 비교	H0 채택	0.918	H0 채택	0.003	0.960
2년간	2001년-2003년 비교	H0 채택	0.198	H0 채택	0.654	0.426
	2003년-2005년 비교	H0 채택	0.371	H0 채택	0.744	0.396
3년간	2002년-2005년 비교	H0 채택	0.290	H0 채택	0.914	0.348
4년간	2001년-2005년 비교	H0 기각 (증가)	0.033	H0 채택	3.646	0.068

위에서 보듯이 전체 평가 평균 검정에서는 2001년부터 2005년을 비교한 4년

차 사이에서만 유의미한 결과가 도출되었고, 분산검정에서는 유의미한 결과가 나오지 않았다. 공기업 평가제도는 단기적으로는 큰 효과가 존재하지 않았으며, 4년의 장기에는 분산의 감소를 이끌어내지는 못했지만 평균적인 성과를 증진시키는 성과가 있었다고 볼 수 있다.

종합 경영 부문의 변화 여부를 검정하면 다음의 <표 2>와 같다.

<표 2> 종합경영 부문의 평균 및 분산 변화 검정

기간	가 설	평균 검정		분산 검정		
		H0:차이없음 H1:차이있음	P 값	H0:차이없음 H1:차이있음	F 값	P 값
1년간	2001년-2002년 비교	H0 기각 (감소)	0.003	H0 기각 (분산 감소)	8.643	0.0071
	2002년-2003년 비교	H0 채택	0.766	H0 채택	0.022	0.882
2년간	2001년-2003년 비교	H0 기각 (감소)	0.005	H0 기각 (분산 감소)	9.718	0.004
	2003년-2005년 비교	H0 기각 (증가)	0.000	H0 기각 (분산 증가)	22.29	8.4E
3년간	2002년-2005년 비교	H0 기각 (증가)	3.9E	H0 기각 (분산 증가)	20.76	0.000
4년간	2001년-2005년 비교	H0 기각 (증가)	0.040	H0 채택	3.471	0.074

종합 경영 부문에서는 평균 검정의 경우 단기 및 장기에 걸쳐서 변화가 있다고 판단할 수 있으며, 분산 검정의 경우 단기에는 변화가 있지만 4년 장기의 경우에는 변화가 없다고 판단할 수 있다. 그런데 이 변화의 세부적인 내용을 살펴보면, 2001년에서 2002년 사이, 그리고 2001년에서 2003년 사이는 해당 부문의 평균이 감소하였고, 나머지 부문은 해당 부문의 평균이 증가한 경우이다. 즉 2001년에서 2003년 사이에는 유의미한 하락이 있었고, 2003년 이후에는 유의미한 상승이 있었다. 따라서 평균 검정에서 단기적으로는 2003년 이후에 개선효과가 있었으며, 장기적으로는 3년 이상의 기간에서 개선 효과가 있다고 판단될 수 있다. 그런데 분산검정을 고려하면, 평균이 상승한 2002-2005년, 2003-2005년 기간 동안 분산은 오히려 증가하였다. 평균이 증가하였지만 분산도 증가하여, 개선 효과는 있지만 그 정도는 크지 않다고 볼 수 있다. 2001-2005년 사이에서는 분산 변화 없는 평균 증가의 개선 효과가 있었다.

다음으로 주요 사업 부문의 변화 여부를 검정하면 다음의 <표 3>과 같다.

<표 3> 주요 사업 부문의 평균 및 분산 변화 검정

기간	가 설	평균 검정		분산 검정		
		H0:차이없음 H1:차이있음	P 값	H0:차이없음 H1:차이있음	F 값	P 값
1년간	2001년-2002년 비교	H0 채택	0.519	H0 채택	0.359	0.559
	2002년-2003년 비교	H0 기각 (감소)	0.015	H0 채택	0.857	0.363
2년간	2001년-2003년 비교	H0 채택	0.571	H0 채택	0.300	0.588
	2003년-2005년 비교	H0 기각 (증가)	0.049	H0 채택	1.513	0.230
3년간	2002년-2005년 비교	H0 채택	0.416	H0 채택	0.228	0.637
4년간	2001년-2005년 비교	H0 채택	0.363	H0 채택	0.969	0.334

주요 사업 부문은 가설이 기각되는 경우가 거의 없었다. 평균 검정의 경우 2002년과 2003년 사이, 그리고 2003년과 2005년 사이에 유의미한 차이가 있었지만, 2002년과 2003년 사이는 평균이 감소한 경우이고, 2003년과 2005년 사이에만 평균이 증가하였다. 그리고 분산 검정의 경우 모든 경우에서 가설이 기각되지 않았다. 2003년-2005년 사이에서 분산 감소가 없는 평균 증가가 있었다고 판단할 수 있다.

마지막으로 경영관리 부문의 변화 여부를 검정하면 다음의 <표 4>와 같다.

<표 4> 경영관리 부문의 평균 및 분산 변화 검정

기간	가 설	평균 검정		분산 검정		
		H0:차이없음 H1:차이있음	P 값	H0:차이없음 H1:차이있음	F 값	P 값
1년간	2001년-2002년 비교	H0 기각 (증가)	0.000	H0 채택	1.177	0.288
	2002년-2003년 비교	H0 채택	0.467	H0 채택	0.060	0.807
2년간	2001년-2003년 비교	H0 기각 (증가)	0.004	H0 채택	1.719	0.202
	2003년-2005년 비교	H0 기각 (감소)	5.5E	H0 기각 (분산 감소)	40.270	1.4E
3년간	2002년-2005년 비교	H0 기각 (감소)	2.4E	H0 기각 (분산 감소)	36.173	3.2E
4년간	2001년-2005년 비교	H0 기각 (감소)	0.000	H0 기각 (분산 감소)	15.950	0.000

경영관리 부문의 결과를 살펴보면, 2001-2002년, 2001-2003년 기간 동안은 분산의 변화가 없는 평균 증가가 있었다. 그러나 2003-2005년, 그리고 3년 내지 4

년의 장기간 동안에는 평균의 감소와 분산의 감소가 있었다. 평균과 분산이 동시에 감소하고 있으므로 경영관리 부문에서는 평가 제도 시행으로 인한 개선 효과가 존재하지 않고, 오히려 해당 부문의 상태가 악화된 것으로 판단된다.

2. 왜도 첨도 분석

앞에서 살펴본 2001년도부터 2005년도까지 각 공기업 평가의 평균, 분산 결과에 왜도 및 첨도를 추가하면 다음과 같다.

〈표 5〉 공기업 전체 평가 결과 평가의 연도별 통계량 변화

구 분	기 간	평균	분산	왜도	첨도
1년간	2001년과 2002년 비교	-	-	↑	↑
	2002년과 2003년 비교	-	-	↑	↓
2년간	2001년과 2003년 비교	-	-	↑	↑
	2003년과 2005년 비교	-	-	↑	↑
3년간	2002년과 2005년 비교	-	-	↑	↑
4년간	2001년과 2005년 비교	↑	-	↑	↑

벽이 있는 분포의 경우, 평균이 동일할 때 왜도가 감소할수록 벽에 더 가까워져서 개선이 있었다고 판단할 수 있다. 그리고 왜도가 감소하거나 동일할 때 첨도가 증가할 경우 개선이 있다고 평가할 수 있다. 그런데 위에서 보듯이, 2001년부터 2005년도까지 전체적으로 왜도는 증가하고 있고, 첨도도 증가하고 있다. 평균하고 분산이 동일한 상태에서 왜도와 첨도가 증가하는 것은, 전체적 평균은 동일하지만 전체적으로 분포가 평균보다 낮은 점수에서 몰리는 형태로 분포 모양이 변화되고 있다는 것을 의미한다. 벽이 있는 상태에서 이러한 분포의 변화는 성과의 쇠퇴를 뜻한다. 그리고 2001-2005년 장기간의 경우 평균은 상승하지만 왜도와 첨도가 같이 증가하고 있으므로, 그 성과 개선의 정도는 그만큼 악화된다고 볼 수 있다.

종합 경영 부문의 평균, 분산, 왜도, 첨도의 변화를 살펴보면 다음과 같다.

〈표 6〉 종합경영 부문의 연도별 통계량 변화

구 분	기 간	평균	분산	왜도	첨도
1년간	2001년과 2002년 비교	↓	↓	↑	↓
	2002년과 2003년 비교	-	-	↑	↑
2년간	2001년과 2003년 비교	↓	↓	↑	↑
	2003년과 2005년 비교	↑	↑	↓	↑
3년간	2002년과 2005년 비교	↑	↑	↑	↑
4년간	2001년과 2005년 비교	↑	-	↑	↑

종합 경영 부문에서 평균과 분산만 고려할 경우 2002-2003년은 정체 상태였고, 2003-2005년, 2002-2005년은 약한 개선의 효과가 인정되며, 2001-2005년은 개선이라고 인정될 수 있었다. 그런데 왜도와 첨도를 같이 고려할 경우, 2002-2003년은 평균과 분산은 변화가 없지만 왜도와 첨도가 증가하였다. 분포가 100점 만점 벽에서 멀어지는 방향으로 이동하였다. 2003-2005년은 평균과 분산을 볼 때 약한 개선이었지만 왜도가 낮아지고 첨도가 높아져서 개선의 효과가 상대적으로 증가했다고 볼 수 있다. 그리고 2002-2005년, 2001-2005년은 왜도와 첨도가 같이 증가하여 개선의 효과가 더욱 낮아지고 있다.

주요 사업 부문의 평균, 분산, 왜도, 첨도의 변화를 살펴보면 다음과 같다.

〈표 7〉 주요사업 부문의 연도별 통계량 변화

구 분	기 간	평균	분산	왜도	첨도
1년간	2001년과 2002년 비교	-	-	↑	↓
	2002년과 2003년 비교	↓	-	↓	↑
2년간	2001년과 2003년 비교	-	-	↑	↓
	2003년과 2005년 비교	↑	-	↓	↑
3년간	2002년과 2005년 비교	-	-	↓	↑
4년간	2001년과 2005년 비교	-	-	↑	↓

주요 사업 부문은 평균과 분산만 고려할 경우에는 2002년-2003년간은 쇠퇴, 2003-2005년간은 개선이 인정되고 나머지 부분은 정체였다. 그러나 왜도와 첨도를 같이 고려할 경우 2002-2005년간은 평균 변화는 없지만 왜도가 하락하면서 첨도가 상승하여 개선 효과를 인정할 수 있다. 그리고 2001-2002, 2001-2003, 2001-2005년간은 평균과 분산의 변화는 없지만 왜도가 상승하고 첨도는 하락하여 전반적으로 낮은 쪽으로 분포가 이동하여 성과의 쇠퇴를 보여준다. 2002년이

나 2003년에서 2005년까지 중기 동안은 성과 향상이 보여지지만, 단기간이나 장기간에 걸쳐서는 성과 개선이 이루어지지 못하고 하락 경향을 보여준다.

경영관리 부문의 평균, 분산, 왜도, 첨도의 변화를 살펴보면 다음과 같다.

<표 8> 경영관리 부문의 연도별 통계량 변화

구분	기간	평균	분산	왜도	첨도
1년간	2001년과 2002년 비교	↑	-	↑	↑
	2002년과 2003년 비교	-	-	↑	↓
2년간	2001년과 2003년 비교	↑	-	↑	↑
	2003년과 2005년 비교	↓	↓	↓	↓
3년간	2002년과 2005년 비교	↓	↓	↑	↓
4년간	2001년과 2005년 비교	↓	↓	↑	↓

평균과 분산만을 볼 때 제도가 시행된 단기간 동안에는 성과 개선이 이루어졌으나, 후반기에는 성과 하락이 발생하였었다. 왜도와 첨도를 고려하면, 평균이 상승한 초반기에는 왜도와 첨도가 같이 상승하고 있다. 즉 평균은 상승하지만 분포는 낮은 점수대로 모여지는 형태를 보여 성과 향상의 정도가 평균만 고려할 경우보다 낮다고 볼 수 있다. 그리고 성과 하락이 발생하는 후반기에는 왜도가 상승하고 첨도는 하락하여, 평균과 분산만 고려할 때보다 하락의 정도가 심하다는 것을 알 수 있다.

이상에서 살펴본 공기업 평가 각 부문의 평균, 분산, 왜도, 첨도를 모두 종합한 결과는 다음과 같다.

<표 9> 평가 종합

구분	기간	전체점수					종합경영					주요사업					경영관리					
		평균	분산	왜도	첨도	평가	평균	분산	왜도	첨도	평가	평균	분산	왜도	첨도	평가	평균	분산	왜도	첨도	평가	
1년간	01-02	-	-	↑	↑	↓	↓	↓	↑	↓	↓	-	-	↑	↓	↓	↑	-	↑	↑	↑	약
	02-03	-	-	↑	↓	↓	-	-	↑	↑	↓	↓	-	↓	↑	↓	-	-	↑	↓	↓	↓
2년간	01-03	-	-	↑	↑	↓	↓	↓	↑	↑	↓	-	-	↑	↓	↓	↑	-	↑	↑	↑	약
	03-05	-	-	↑	↑	↓	↑	↑	↓	↑	↑	↑	-	↓	↑	↑	↓	↓	↓	↓	↓	↓
3년간	02-05	-	-	↑	↑	↓	↑	↑	↑	↑	↑	-	-	↓	↑	↑	↓	↓	↑	↓	↓	↓
4년간	01-05	↑	-	↑	↑	↑	↑	-	↑	↑	↑	-	-	↑	↓	↓	↓	↓	↑	↓	↓	↓

전체적으로 볼 때 공기업 평가는 단기간 및 중간 기간 동안 성과 개선을 이

끝지 못했다. 단 2001년에서 2005년의 장기간 동안에는 약간의 성과 개선 효과가 존재하였다. 각 부문별로 살펴보면 종합경영 부문은 단기간 동안에는 성과 개선이 이루어지지 못하였으나, 장기간에는 약간의 성과 개선이 발생하였다. 주요 사업 부문은 단기적 및 장기적으로 평균은 정체 상태이나 왜도 상승, 첨도 하락으로 약간의 성과 하락이라고 볼 수 있다. 그리고 경영관리 부문은 단기간 동안에는 약간 개선 효과가 있었으나 장기간 동안 평균, 분산이 감소하고 왜도 상승, 첨도 하락으로 성과 하락의 폭이 컸다. 결국 단기적으로는 종합경영, 주요 사업, 전체 종합 부문이 성과 하락을 보이고 있으며, 장기적으로는 사업 부문, 경영관리 부문은 하락하였으나 종합경영 부문이 상승하여 전체적으로 약간의 상승효과를 보이고 있다.

V. 결 론

지금까지 최근의 수개년 동안의 공기업 경영평가 결과를 바탕으로, 해당 평가 부문에 대한 실질적인 개선 효과가 존재하는가를 살펴보았다. 종합 결과를 볼 때 평균과 분산만 고려할 경우 공기업 평가는 단기적으로 큰 효과가 없고 장기적으로는 향상의 효과가 인정되었으나, 왜도와 첨도를 고려할 경우 단기적으로는 성과 하락이 이루어지고, 장기적으로도 향상의 효과가 적었다. 그리고 주요 사업 부문의 경우에도 평균과 분산만 고려할 경우에는 큰 변화가 없었으나, 왜도와 첨도를 같이 고려할 경우 성과 하락이 인정될 수 있었다. 종합 경영의 경우 왜도와 첨도를 고려하면 성과 향상의 정도가 낮아졌으며, 경영관리의 경우 왜도와 첨도를 고려할 경우 성과 하락의 정도가 더욱 심해졌다.

이러한 분석 결과를 볼 때, 현재 공기업 평가 제도는 비록 해마다 평가를 하고 순위를 매기기는 하지만, 전반적으로 평가의 목적인 공기업 평가 부문의 향상 및 개선에는 큰 효과를 내지 못하고 있다고 볼 수 있다.

본 연구의 정책적 함의는 다음과 같다. 첫째, 현재 공기업 평가 지표는 진정한 공기업의 기업 개선 노력을 반영하지 못할 수 있으므로, 평가 지표에 대한 검토가 요구된다. 본 평가 결과는 전반적으로 상승 추세를 보여주지 못하지만, 공기업의 효율성 및 생산성을 연구하는 여타 많은 연구들은 공기업의 효율성이 지속적으로 증가하고 있다는 것을 보여주고 있다. 공기업의 실제 생산성 및 효율성이 계속 상승함에도 불구하고 평가 지표 결과가 향상되지 않고 있다는 것은 평가 지표가 공기업의 생산성과 효율성을 제대로 반영하지 못하고 있다는 것을 의미한다. 따라서 각 평가 지표가 보다 공기업의 효율성 및 생산성과 직접적인 연관을 맺을 수 있도록 지표에 대한 조정이 요구된다.

〈표 10〉 평균 분산 검증과 왜도 첨도 분석 결과

기간		전체점수		종합경영		주요사업		경영관리	
		평균 분산	왜도 첨도 추가	평균 분산	왜도 첨도 추가	평균 분산	왜도 첨도 추가	평균 분산	왜도 첨도 추가
1년간	2001년과 2002년 비교	-	후퇴	후퇴	후퇴	-	후퇴	개선	개선 (약)
	2002년과 2003년 비교	-	후퇴	-	후퇴	후퇴	후퇴	-	후퇴
2년간	2001년과 2003년 비교	-	후퇴	후퇴	후퇴	-	후퇴	개선	개선 (약)
	2003년과 2005년 비교	-	후퇴	개선 (약)	개선	개선	개선	후퇴	후퇴
3년간	2002년과 2005년 비교	-	후퇴	개선 (약)	개선 (약)	-	후퇴	후퇴	후퇴 (강)
4년간	2001년과 2005년 비교	개선	개선 (약)	개선	개선 (약)	-	후퇴	후퇴	후퇴 (강)

둘째, 공기업 평가는 상대적인 평가 결과가 아니라 절대적인 평가를 중시하는 방향으로 보완이 요구된다. 위와 같이 공기업 평가 제도가 해당 부문에서 점수 향상이 제대로 이루어지지 못한 이유 중 한 가지는 현재 공기업 평가에 대해서 각 공기업이 보다 중요하게 생각하는 것은 절대적 점수가 아니라 다른 공기업들에 대비한 상대적 점수이기 때문인 것으로 볼 수 있다. 평가 점수에 의해서 각 공기업의 순위가 매겨지는 상황에서, 각 공기업들은 각 평가 부문에서 절대적 점수를 높게 받는 것보다는 다른 공기업보다 높은 점수를 받는 상대적 점수에 더 관심을 가질 가능성이 크다. 비록 매해 평가가 이루어진다 하더라도 각 평가 부문에서 절대적 수준을 높여야 할 필연적인 이유는 없으며, 따라서 평가가 수년간 이루어졌다 하더라도 해당 부문이 실질적으로 향상되지 않을 수 있다. 이러한 평가의 오류를 개선하기 위해서는 절대적인 개선도를 중시할 필요성이 존재한다. 공기업 평가가 단순히 평가대상 공기업들의 상대적 순위만을 중시하는 것이 아니라 절대적 수준 상승을 추구할 수 있도록 제도적 기반을 마련할 필요가 있다.

셋째, 평가지표 해석 도구로서 평균, 분산 외에 왜도와 첨도를 활용하는 것을 고려할 수 있다. 일반적으로는 평균을 가지고 평가를 해석하고 있는바, 위에서 본 바와 같이 왜도와 첨도를 활용할 경우 평균 점수가 제시하는 평가 결과보다 더 다양한 평가 측면을 제시해줄 수 있다. 또한 왜도와 첨도는 해당 평가 부문이 아직 개선의 여지가 많이 남아 있는지, 아니면 평가대상 기관들이 전반적으로 높은 수준을 달성하여 더 이상 개선의 여지가 많이 남아있지 않은지 하는가에 대한 지표를 제공해줄 수 있다. 점수의 벽에 부딪혀서 더 이상 개선의 여지

가 많이 남아있지 않은 부문을 대상으로 지속적으로 똑같이 평가를 하는 것은 평가의 효과를 저하시킬 수 있다. 어느 정도 수준에 도달한 지표는 그 비중을 감소시키고 아직 부족한 부문에 대한 지표를 추가하는 것이 평가로 인한 개선의 효과를 극대화시킬 수 있는바, 왜도와 첨도 분석은 그러한 지표 개선의 필요성을 제시해주는 역할을 할 수 있다.

그리고 본 연구의 한계는 다음과 같다. 첫째, 본 연구에서는 공기업의 비계량적 평가 점수를 연구 대상으로 하였는바, 비계량 평가 점수는 평가자의 주관성을 완전히 배제할 수 없는 한계가 있다. 비계량평가 점수는 매년 평가단의 인적 구성이 변화하는데서 발생하는 점수 편차가 존재하고, 또한 후광효과 등 평가자 개인의 무의식적인 평가오류가 발생할 가능성이 크다. 이와 같이 평가자의 주관성 배제에 한계가 있는 비계량적 지표 점수를 활용하여 평균, 분산, 왜도, 첨도 분석을 하는 것은 연구결과와 객관성에도 영향을 미칠 수 있다.

둘째, 본 연구의 대상은 2001년부터 2005년까지로 전체 공기업 평가 기간을 고려할 때 상당히 짧은 시간만을 대상으로 하고 있다. 공기업 평가 제도는 1980년대부터 시행되어 왔는바, 본 연구는 공기업 평가제도 전반에 대한 검토로는 한계가 존재한다. 이러한 한계가 존재하지만, 본 분석은 평균, 분산, 왜도, 첨도의 분석 방법으로 공기업 평가 제도를 평가하고자 한 것으로, 이러한 방법을 활용하여 공기업 평가제도를 분석하고자 하는 범위 내에서 본 연구는 어느 정도 의미를 지닐 수 있을 것이다.

참고문헌

- 공병천. (2005). 상위평가에 대한 비판적 논고: 중앙행정기관 평가를 중심으로. 『한국행정연구』 14(1).
- 곽병한. (2005). 정부투자기관 경영평가지표의 인과관계에 관한 실증분석. 한국유통과 학회 2005년 동계 정기학술대회.
- 곽채기. (2003). 정부투자기관의 비효율성 통제를 위한 경영평가제도의 역할 및 운영 성과. 『공기업논총』 15(1).
- 기획예산처. 각연도 정부투자기관 경영평가 보고서.
- 김명수. (1987). 『공공정책평가론』. 박영사.
- 김양렬·민충기. (2005). 『경영경제 통계학』. 명경사.
- 김종희. (2006). 지방공기업 경영평가의 서열화와 비계량지표에 관한 고찰. 『한국지방공기업학보』 2(1).
- 김준기. (2001). 공기업경영평가제도의 성과와 향후 과제. 『한국행정연구』 10(1).

pp.97-12.

- 나 영·이창우. (2001). 공기업 경영평가시스템 운용의 유용성에 관한 실증연구. 『회계학연구』 26(2).
- 노화준. (1995). 『정책평가론』. 법문사.
- _____. (2001). 공공부문 성과측정의 이슈와 정책평가제도의 개혁방안. 서울대 행정대학원 BK21 교수 논문집.
- 문영세. (2006). 한국 중앙심사평가제도의 메타평가적 분석. 한국행정학회 동계학술대회 발표문.
- 문원식. (2005). 지방정부 혁신평가에 대한 비교연구. 『지방행정연구』 19(4):33-59.
- 박명호·이성욱. (2000). 정부투자기관 경영평가를 위한 비계량 평가지표의 개선방향. 『공기업 논총』 12(1).
- 박석희. (2003). 정부투자기관 경영평가제도 20년의 조망과 변천과정 분석. 『공기업논총』 15(1) pp.11-47.
- _____. (2006). 공공부문의 성과관리와 조직역량: 차단된 시계열설계를 통한 13개 공기업의 실증분석. 『한국행정학보』 40(3). pp.219-244.
- 박중수. (2003). 정보화사업 메타평가를 위한 CIPOU 모형개발. 한국정책학회 추계학술대회 발표논문.
- 박영범. (1998). 경영환경변화의 수용을 위한 정부투자기관 경영평가제도의 비계량지표의 개선방안. 『공기업논총』 10(1).
- 배용수. (2006). 지방공기업 경영혁신, 어디까지 왔나 ; 지방공기업 경영평가제도의 개선방향. 『지방재정』 9(5). 한국지방재정공제회.
- 백승천. (2003). 지방공기업 경영평가지표 선정 및 가중치 설정. 고려대학교 박사학위 논문.
- 성태제. (2003). 감사나 평가활동에 대한 메타평가적 관점에서의 결과타당도. 『교육학연구』 41(1) pp.91-110.
- 송대회·송명희(1994). 우리나라 공기업 책임 경영제도 운영에 대한 평가 및 정책방향. 『공기업 논총』 6(1). 한국공기업학회.
- 송대회. (2003). 공기업 경영평가제도의 회고와 발전 방향. 『공기업논총』 15(1) pp. 1-10.
- 신 열. (2002). 지방공기업 경영평가제도의 전략적 발전방안: SER-M SWOT 방법의 활용. 『한국정책학회보』 11(4).
- _____. (2004). 지방공기업 경영평가제도에 대한 재검토. 『한국행정연구』 13(3).
- 신 열·오철호. (2002). 지방공기업 경영평가제도의 비판적 재검토: 평가 체제, 지표 및 방법을 중심으로. 『지역정책연구』 13(2). 충북개발연구원.
- 오연천. (1996). 경영평가제도의 회고와 방향. 『행정논총』 35(1).

- _____. (1998). 공기업 구조조정과 경영평가제도의 개선 방향. 『공기업논총』 10(1). pp.3-10.
- 오연천·곽채기. (2003). 정부투자기관 경영평가제도의 개선 방안. 『공기업논총』 15(1). pp.123-154.
- 이규환. (2000). 우리나라 지방공기업의 발전방향. 『중앙행정논집』 14(1). 중앙대학교 국가정책연구소.
- 이상철·성도경(2002). 지방공기업의 책임경영을 위한 성공요건 분석: 뉴거버넌스의 관점에서. 『한국행정학보』 36(4).
- 이석원. (2005). 조정성과지표를 사용한 공공기관의 성과측정 : 정부투자기관경영평가에의 적용.
- 이종수. (2003). 지방공기업 경영평가지표 개발에 관한 연구: 조직문화지표를 중심으로. 『중앙행정논집』 17(1). 중앙대학교 국가정책연구소.
- 이형우·임동환. (2005). BSC의 관점에서 공기업의 성과평가 지표 개선에 관한 연구. 『부동산 학보』 25.
- 이환범·송건섭·김병문. (2005). 지방공기업의 경영성과관리와 평가지표개발 : 자료포락 분석(DEA)에 의한 효율성 측정. 『한국사회와 행정연구』 15(4).
- 임성욱. (2003). 중앙정부의 사회복지시설 평가에 대한 비판적 분석 -메타평가를 중심으로-. 『정부학연구』 9(2). pp.261-266.
- 이종욱. (2003). 기금의 평가와 성과개선. 『경제학연구』 51(4).
- 이찬구. (2003). 영국의 범부처간 과학기술정책 평가제도-메타평가와 한국에의 시사점 -. 『한국행정논집』 15(4). pp.809-833.
- 장지인·이정구(1998). 공기업 경영혁신을 위한 계량평가지표의 개선방향. 『공기업 논총』 10(1).
- 장지인·곽채기. (2000). 정부투자기관 경영평가제도 및 운용체계의 개편 방안. 『공기업 논총』 12(1). pp.39-83.
- 황병상·강근복. (2005). 정부출연 연구기관의 평가의 발전방안 논고: 기초기술연구회의 평가 사례에 대한 메타평가를 중심으로. 『한국정책학회보』 14(1).
- 황혜신. (2006). 메타방법론으로서의 Fischer의 정책평가의 논리에 관한 고찰 -국민연금 사례에의 적용을 중심으로-. 『한국행정논집』 18(4). pp.1179-1204.
- Chelimsky, Eleanor ed.. (1985). *Program Education: Patton and Directions*. Washington, D.C. :The American Society for Public Administration
- Cook, Thomas D and Gruder, Charles L.(1978). "Metaevaluation" *Evaluation Quarterly*. Vol.2. No.1. 5-51
- Cordray, D.S. Fisher, R.L.(1994). Synthesizing evaluation findings. In. Wholey, J. S.,

- Hatry, H.P. &Newcomer, K. E.(Ed).(2004). The Handbook of Practical Program Evaluation . San Francisco: Jossey-Bass.
- GAO(1992). *The Evaluating Synthesis*.
- Gould, Stephen Jay. (1996). *Full House*. Harmony.
- Larson, Richard and Berliner, Leni. (1983). "On Evaluating Evaluation". *Policy Sciences*. Vol.16, No.2. 147-163.
- Nilsson, N. & Hogben, D. (1983). "Metaevaluation" in house, E. R.(ed). *Philosophy of Evaluation* . San Francisco :Jossey-Bass. pp.83-97.
- Ogata, P.T.,(1940)."Evaluating Evaluation". *Journal of Educational Research*. Vol.33.
- Schneider, Anne Larason. (1986). "The Evaluation of a Policy Orientation for Evaluation Research: A Guide to Practice", *Public Administration Review*, Vol.46, No.4.
- Scriven, Michael. (1969). "An Introduction to Meta-evaluation". *Educational Product Report*. Vol.2, 36-38
- Scriven M. (1991). *Evaluation Thesaurus* (4th de). Newbery Park. CA, Sage Publications.
- Stufflebeam L. (1974). Meta-evaluation (Occasional Paper No 3). The Evaluating Center, Western Michigan University.
- Stufflebeam, Daniel. (1981). "Metaevaluation: Concept, Standard, and Uses", in Ronald A.Berk. ed. *Educational Evaluation Methodology*. The Johnes Hopkins University Press.

Abstract

Performance Analysis of the Public Enterprise Evaluation System

SeongRak Choi·MinJeong Park

This study examines the public enterprise evaluation system, which is one of the most representative evaluation systems. The research question is whether the public enterprise evaluation system improved evaluation criteria, as initially expected, by using the method of descriptive statistics and testing hypotheses.

The results show that no meaningful improvement has been made to any of the criteria of the public enterprise evaluation system. Other studies show that the overall efficiency of public enterprise is improving but there is no improvement within the public enterprise evaluation system. This suggests that the current public enterprise evaluation system does not understand the efficiency of public enterprise and that there are problems with the evaluation system and the criteria.

【Key words: public enterprise, public enterprise evaluation, meta evaluation】